UNITED STATES DEPARTMENT OF COMMERCE
Bureau of the Census
Washington, DC 20233-0001

September 22, 1999

DSSD CENSUS 2000 PROCEDURES AND OPERATIONS MEMORANDUM SERIES R-19

MEMORANDUM FOR Howard Hogan
Chief, Decennial Statistical Studies Division

From: Donna Kostanich DK
Assistant Division Chief, Sampling and Estimation
Decennial Statistical Studies Division

Prepared by: Thomas Mule TM
Sample Design Team

Subject: Accuracy and Coverage Evaluation Survey: Differential Allocation of Sample within State by Demographics

I. INTRODUCTION

The Accuracy and Coverage Evaluation (A.C.E.) survey interview sample of 300,000 housing units has been allocated to the 50 states and the District of Columbia as given in Mule (June, 1999). The next step is to determine how to allocate the sample within each state by demographic groups. Differential sampling is being investigated to increase sample sizes for the historically undercounted groups. A good differential sampling plan should improve reliability of the smaller population subgroups while having minimal impact on other subgroups. In addition to improving reliability, increasing sample size of small population groups is important for poststratification plans.

The A.C.E. survey design plan involves several stages of selection. First, approximately 30,000 block clusters have been selected using the original ICM sampling plan. This sample is about 2.5 times larger than planned for the A.C.E. survey. This sample is called the A.C.E. listing sample. Field staff will visit each cluster in the A.C.E. listing sample and will create an independent address list. Then, the A.C.E. listing sample will be reduced using the differential sampling plan developed based in part from this research. Two additional subsampling operations will be done: small block subsampling and large block subsampling.

The within-state allocation by demographic groups is one component of an A.C.E. differential sampling plan. Although not discussed in this memorandum, the other component is differential

sampling by consistency between the count of housing units from the census and the independent list. More information on this research is in Farber (September, 1999).

II. A.C.E. LISTING SAMPLE

The basis of our research is the A.C.E. listing sample. This is the originally planned ICM design with some minor state allocation modifications. The A.C.E. listing sample clusters have been selected.

A. SAMPLE DESIGN

The A.C.E. listing sample consists of almost 30,000 block clusters. A block cluster is a group of one or more contiguous collection blocks. Block clusters were defined for the 50 states, the District of Columbia and Puerto Rico but not for remote areas of Alaska.

Before the listing sample was selected, each block cluster was classified into one of 12 demographic/tenure groups based on the estimated 1990 race/ethnicity and tenure composition of the cluster. The six race/ethnicity categories are American Indian, Asian, Hawaiian and Pacific Islander, Hispanic, Black and White and Other. Each of these six categories is crossed by tenure to yield the 12 demographic/tenure groups. This variable was used as a sort variable during the A.C.E. listing sample selection. Each cluster was classified into an estimated 1990 urbanized group: urban 250,000+, other urban, or non-urban.

Each cluster was assigned to one of four sampling strata in each state:

- Small block clusters
- Medium block clusters
- Large block clusters
- American Indian Reservations

All block clusters with 0 to 2 housing units were in the small block cluster stratum. Block clusters with 3 or more housing units and on American Indian Reservations were in the American Indian Reservations stratum. See Cromar (May, 1999) for exceptions. For the remaining clusters, 3 to 79 housing units are medium and 80 or more housing units are large.

For each state, a systematic sample of block clusters was drawn from each of the sampling strata at variable rates. Large block clusters were sampled at a higher rate than medium block clusters since a later subsampling operation will decrease the overall sampling rate of housing units in these clusters.

For more information on the listing sample selection, see Salganik and ZuWallack (March, 1999) and Mule (March, 1999).

B. CLUSTER DATA

For each cluster in the listing sample, we have an estimate of the expected number of housing units based on an early version of the census address list, and an expected number of persons by demographic and tenure composition. See Appendix A for further documentation of these cluster characteristics and expected sample sizes. Using this information, we can develop within-state allocations and simulate the expected reliability for estimates of interest.

In order to simulate the reliability of the various sample designs, large block cluster subsampling needs to be reflected. If a cluster with over 80 housing units is selected in the A.C.E. reduction sample, only a subsample of those housing units would actually be interviewed in the A.C.E. survey. For this research, the large block cluster subsampling plan was simulated within each state by generating an expected sample size when there is equal weighting for the medium and large block clusters.

III. ALTERNATIVE DESIGNS

A. REDUCTION SUBSAMPLING STRATA

For each state, reduction subsampling strata will need to be determined for the A.C.E. reduction. These subsampling strata can be sampled at different rates in order to increase sample sizes of small population groups and improve the reliability of some estimates. For the listing sample, each cluster was classified into 1 of 12 demographic/tenure categories. Table 1 below shows the classification.

Table 1
Cluster Demographic/Tenure Classification

Order	Criteria
1	Proportion of Hawaiian and Pacific Islander Renters ≥ 0.10
2	Proportion of Hawaiian and Pacific Islander Owners ≥ 0.10
3	Proportion of American Indian and Alaska Native Renters ≥ 0.10
4	Proportion of American Indian and Alaska Native Owners ≥ 0.10
5	Proportion of Asian Renters ≥ 0.20
6	Proportion of Asian Owners ≥ 0.20
7	Proportion of Hispanic Renters ≥ 0.20
8	Proportion of Hispanic Owners ≥ 0.20
9	Proportion of Black Renters ≥ 0.25
10	Proportion of Black Owners ≥ 0.25
11	Proportion of Other Renters ≥ 0.30
12	all else

The first subsampling strata option involves forming two reduction strata in each state. Clusters will be considered minority based on their Demographic/Tenure group code. This will be referred to as the Demographic/Tenure Group subsampling strata plan. Table 2 shows this classification.

Table 2
Demographic/Tenure Group Subsampling Strata Plan Classification

Order Number	Demographic/Tenure Group Subsampling Stratum Plan
1 - 10	Minority
11 - 12	Non-Minority

A concern with the Demographic/Tenure Group plan is that the proportion of a single race/ethnicity and tenure group is relatively low, yet the cluster is being classified as a minority. Another concern is that a cluster could have people in multiple race groups yet not be classified as minority because of the inclusion of tenure. Therefore, another alternative is considered.

The second subsampling strata option involves classifying clusters as minority based on the percentage of minorities that lived in the cluster in 1990. This will be referred to as the Minority 50% Plus subsampling strata plan. Table 3 shows this classification.

Table 3
Minority 50% Plus Subsampling Strata Classification

Percent Minority in 1990	Minority 50% Plus Subsampling Stratum Plan
Greater than or equal to 50%	Minority
Less than 50%	Non-Minority

B. ALLOCATIONS

This section lists the allocations examined in this preliminary research.

1. PROPORTIONAL ALLOCATION

Proportional allocation of sample within a state minimizes the amount of differential weighting that is introduced by the sample design. This allocation will serve as a benchmark for comparing differential subsampling allocations.

2. OPTIMIZING 1990 MAJOR GROUPS

For the two reduction subsampling strata options, we will need to determine the sampling rate for each stratum. We would do differential sampling in order to increase the sample size and improve the reliability of some of our estimates. Since the final A.C.E. poststrata have yet to be determined, the 1990 major group simulated reliability estimates will be generated for 53 major groups. The major groups are the 51 major poststrata from the 1990 PES (357 collapsed across age/sex levels) with separate estimates for the Asians and the Hawaiian and Pacific Islanders. Appendix B discusses an optimal within-state allocation using these 1990 major groups. This allocation will be constrained by the A.C.E. state interview housing unit sample size that was already determined.

The determined allocation is optimal for the 1990 major groups using 1990 PES data. To protect against any shifts in population that may have occurred in the past 10 years, the optimal allocation results will be used with other information as a guide for determining A.C.E. within-state allocations. This information can be used to identify states that do not need to be differentially sampled.

The following three allocations optimizing the 1990 major groups were investigated:

a. The first allocation used the Demographic/Tenure group subsampling strata plan and placed the same emphasis on each of the 53 groups during the optimization. This will be referred to as the Optimal Demographic/Tenure Group allocation.

b. The second allocation used the Minority 50% Plus subsampling strata plan and placed the same emphasis on each of the 53 groups during the optimization. This will be referred to as the Optimal Minority 50% Plus allocation.

c. The third allocation used the Demographic/Tenure Group subsampling strata plan and placed twice as much emphasis on the Minority groups than the Non-Minority groups during the optimization. This will be the Emphasize Minority allocation.

3. KEEP ALL MINORITY CLUSTERS

This allocation attempted to retain all Minority clusters (as classified by the Demographic/Tenure Group subsampling strata plan) in the A.C.E. Reduction. However, retaining all Minority clusters could not be done in some states because the total state interview allocation would have been exceeded. Also in some states, retaining all Minority clusters would have left a very small number of Non-Minority clusters in sample and thus introducing a large amount of weight variation in the state. In states that fell under these two categories, the Minority sample was restricted to the Optimal Demographic/Tenure Group Allocation and the Non-Minority weights were not allowed to be larger than 1000. This will be referred to as the Keep All Minority allocation.

IV. SIMULATING RELIABILITY ESTIMATES

This study will examine simulated coefficients of variation (CVs) for various sample designs. The simulated CVs can show how various designs affect the reliability of certain estimates. The simulated CVs are in expectation and may vary based on the actual sample selected.

The general form for simulating CVs for one of the major groups for a particular allocation is given below. See Appendix C for a more thorough description.

$$\hat{CV}_{A.C.E.} = \sqrt{\frac{\sum_{s,j} n_{A.C.E.,s,j} W^2_{A.C.E.,s,j}}{\sum_{i}^{n_{PES}} W^2_{PES,i}}}\ CV_{PES}$$

where $n_{A.C.E.s,j}$ is the resulting estimated person sample size for the major group from the *j*th subsampling stratum in the *s*th state for the particular allocation,
$W_{A.C.E.,s,j}$ is the simulated weight for the people sampled from *j*th subsampling stratum in the *s*th state for the particular allocation,
n_{PES} is the E-sample person sample size for the major group from the 1990 PES,
$W_{PES,i}$ is the 1990 E-sample final weight for the *i*th person and
CV_{PES} = 1990 PES Major Group CV.

V. PRELIMINARY RESULTS

Preliminary results of the five allocations are given in Attachments 1 through 3. Attachment 1 shows the state differential sampling factor for each design. A differential sampling factor equal to 2 indicates that the Minority stratum is being sampled at twice the rate as the Non-Minority stratum. For the Proportional Allocation, the rate is 1 for every state. For the Optimal Demographic/Tenure Group Allocation, 11 states have the Minority stratum being differentially sampled by more than twice the rate as the Non-Minority stratum. The Optimal Minority 50% Plus and the Emphasize Minority Allocations are similar to the Optimal Demographic/Tenure Group Allocation with 10 and 14 states, respectively. The results for the Keep All Minority Allocation show a huge increase in weight variation between the Minority and Non-Minority strata. Future research will examine having a maximum differential sampling factor equal to 2 in all states to prevent an increase in variance due to weight variation.

Attachment 2 shows the expected simulated CVs for the 1990 major groups. Also, CVs are given for American Indians on American Indian Country and American Indians on American Indian Reservations[1]. The Proportional Allocation shows gains in reliability as compared to the 1990 PES for the Non-Hispanic White and Other groups. The results are mixed for the Minority groups when comparing between the two. The table shows the change in simulated CVs for the four differential within-state sampling allocations. The Optimal Demographic/Tenure Group allocation shows reliability gains for Minority groups especially in the Northeast. The Optimal Minority 50% Plus allocation shows slightly less gains for the Minority groups as compared to

[1] American Indians on American Indian Country would include Indians living on American Indian Reservations and associated Trustlands, Tribal Jurisdiction Statistical Areas, Tribal Designated Statistical Areas and Alaska Native Village Statistical Areas. American Indians on American Indian Reservations include Reservations and associated Trustlands.

the Optimal Demographic/Tenure group allocation. The Emphasizing Minority and the Keep All Minority allocations show only very slight improvements for most of the Minority groups over the Optimal Demographic/Tenure Group allocation. Comparing the CVs for the two definitions of American Indians shows that including American Indians who live on American Indian Country but who do not live on Reservations in the American Indian poststratum could potentially double the CV. This is because of the huge weight variation introduced even though the sample size is increased by roughly 5 percent.

Attachment 3 shows the expected sample size based on simulating the allocation alternatives for each major group. The Non-Hispanic White and Other groups have expected simulated sample sizes twice the size on average for the Proportional Allocation as compared to the 1990 PES E-Sample. For the Minority groups using the Proportional Allocation, there are eight groups that are expected to have less sample than in 1990. The differential within-state sampling shows increases in expected sample sizes for the minority groups. However, there are some minority groups like Black Owners in the Northeast and South Urban 250,000+ areas that have expected sample sizes less than 1990 E-Sample even under the Keep All Minority allocation. The Optimal Minority 50% Plus allocation has less of a sample size increase for minority groups as compared to the Optimal Demographic/Tenure Group allocation. The Keep All Minority allocation does increase the sample size for some Minority Groups as compared to the optimal allocations. Looking back at Attachment 2, there was not a large increase in reliability based on this increase in sample size. The reliability gains from increased sample size are being offset by the differential weighting introduced.

VI. CONCLUSIONS AND FUTURE WORK

Overall, when comparing differential sampling to proportional sampling, we see some gains in reliability and sample size for small population subgroups with a negligible impact on larger subgroups. Except for the Keep All Minority option, there isn't much difference among the differential sampling allocations. The drawback to the Keep All Minority allocation is the amount of differential sampling introduced.

These options use differential sampling in all states. Future work will investigate whether a combination of differential sampling in some states and proportional allocation in the remaining will yield similar results. Also, we will explore possibly limiting the differential sampling factors to avoid high weight variation.

VII. REFERENCES

Cromar (March, 1999), "Accuracy and Coverage Evaluation Survey: American Indian Reservations Sample Design," DSSD CENSUS 2000 PROCEDURES AND OPERATIONS MEMORANDUM SERIES R, Internal Census Bureau Memorandum.

Farber (September, 1999), "Accuracy and Coverage Evaluation Survey: Sample Reduction Overview," DSSD CENSUS 2000 PROCEDURES AND OPERATIONS MEMORANDUM SERIES R, Internal Census Bureau Memorandum.

Mule (March, 1999), "Accuracy and Coverage Evaluation Survey: Block Cluster Sample Selection Specifications," DSSD CENSUS 2000 PROCEDURES AND OPERATIONS MEMORANDUM SERIES R, Internal Census Bureau Memorandum.

Mule (June, 1999), "Accuracy and Coverage Evaluation Survey: State Interviewing Sample Size Estimates," DSSD CENSUS 2000 PROCEDURES AND OPERATIONS MEMORANDUM SERIES R, Internal Census Bureau Memorandum.

Salganik and ZuWallack (March, 1999) "Accuracy and Coverage Evaluation Survey: Universe File and Block Cluster Sampling Parameter File Specification," DSSD CENSUS 2000 PROCEDURES AND OPERATIONS MEMORANDUM SERIES R, Internal Census Bureau Memorandum.

cc: DSSD Census 2000 Procedures and Operations Memorandum Series Distribution List
Statistical Design Team Leaders Distribution List
Sampling and Estimation Staff

Attachment 1

Differential Sampling Factors

Region	State	Allocation Alternatives				
		Proportional Allocation	Optimal D/T Group	Optimal Min. 50% +	Emphasize Minority	Keep All Minority
NE	Connecticut	1.00	1.81	1.78	2.04	4.67
NE	Maine	1.00	1.00	N/A[1]	1.00	4.56
NE	Massachusetts	1.00	1.83	1.79	1.83	1.83
NE	New Hampshire	1.00	1.00	1.00	1.08	4.88
NE	New Jersey	1.00	2.09	1.73	2.15	2.15
NE	New York	1.00	1.75	1.78	1.98	3.32
NE	Pennsylvania	1.00	1.53	1.52	1.53	1.53
NE	Rhode Island	1.00	1.71	1.80	1.98	1.71
NE	Vermont	1.00	1.00	N/A[1]	0.52	4.58
Midwest	Illinois	1.00	1.06	1.06	1.06	1.06
Midwest	Indiana	1.00	1.48	1.46	1.48	1.48
Midwest	Iowa	1.00	1.88	2.35	1.98	3.18
Midwest	Kansas	1.00	2.17	1.98	2.25	3.75
Midwest	Michigan	1.00	1.08	1.08	1.08	1.08
Midwest	Minnesota	1.00	1.81	1.79	1.81	1.81
Midwest	Missouri	1.00	1.78	1.73	1.78	1.78
Midwest	Nebraska	1.00	1.59	1.67	1.66	5.56
Midwest	North Dakota	1.00	1.97	1.06	2.07	4.48
Midwest	Ohio	1.00	1.12	1.12	1.12	1.12
Midwest	South Dakota	1.00	1.19	1.05	1.23	5.17
Midwest	Wisconsin	1.00	1.53	1.51	1.53	1.53
South	Alabama	1.00	2.46	2.09	2.55	5.66
South	Arkansas	1.00	3.06	2.95	3.25	3.06
South	Delaware	1.00	1.77	1.50	1.81	1.77
South	District of Columbia	1.00	1.61	1.39	1.69	1.61
South	Florida	1.00	1.00	1.00	1.00	1.00
South	Georgia	1.00	1.65	1.58	1.65	1.65
South	Kentucky	1.00	1.72	1.20	1.82	3.49
South	Louisiana	1.00	1.42	1.32	1.43	1.42
South	Maryland	1.00	1.84	1.81	1.92	3.84
South	Mississippi	1.00	2.22	2.06	2.26	2.22
South	North Carolina	1.00	1.71	1.63	1.71	1.71
South	Oklahoma	1.00	2.09	1.32	2.09	2.09
South	South Carolina	1.00	2.32	2.25	2.38	2.32
South	Tennessee	1.00	2.00	2.12	2.13	2.81
South	Texas	1.00	1.43	1.29	1.45	1.43
South	Virginia	1.00	1.73	1.65	1.73	1.73
South	West Virginia	1.00	3.21	3.18	3.56	4.76
West	Alaska	1.00	4.14	5.42	4.28	4.14
West	Alabama	1.00	1.74	2.02	1.79	6.27
West	California	1.00	1.42	1.39	1.46	1.42
West	Colorado	1.00	1.44	1.03	1.50	4.93
West	Hawaii	1.00	1.46	1.36	1.47	1.46
West	Idaho	1.00	2.03	1.21	2.20	7.63
West	Montana	1.00	1.00	N/A[1]	1.00	8.20
West	Nevada	1.00	1.38	1.34	1.43	1.38
West	New Mexico	1.00	1.28	1.35	1.30	1.28
West	Oregon	1.00	1.53	1.83	1.58	2.75
West	Utah	1.00	2.11	2.62	2.18	8.09
West	Washington	1.00	1.72	1.70	1.72	1.72
West	Wyoming	1.00	1.04	1.09	1.09	7.56

[1] N/A = no listing sample Cluster had 50 percent or more minorities in the state.

Simulated Coefficients of Variation[1]

Race/Eth.	Tenure	Geo. Group	1990 PES	Prop. Allocation	Difference from Proportional Allocation: Optimal D/T Group	Optimal Minority 50% +	Emph. Minority	Keep All Minority
		North East:						
NHWO	Owner	Urb>250000	1.06	0.67	0.04	0.02	0.05	0.08
		Other Urban	0.49	0.20	0.00	0.00	0.01	0.02
		Non-Urban	0.69	0.35	0.02	0.01	0.03	0.04
	Renter	Urb>250000	1.41	0.79	0.04	0.02	0.04	0.09
		Other Urban	1.56	1.16	0.04	0.02	0.05	0.10
		Non-Urban	4.37	3.42	0.16	0.09	0.19	0.39
Black	Owner	Urb>250000	1.93	2.24	-0.29	-0.24	-0.30	-0.33
	Renter	Urb>250000	1.76	1.71	-0.25	-0.18	-0.27	-0.33
NBHIS	Owner	Urb>250000	4.41	3.60	-0.12	-0.05	-0.12	-0.03
	Renter	Urb>250000	3.77	2.54	-0.30	-0.15	-0.34	-0.40
		Midwest:						
NHWO	Owner	Urb>250000	0.39	0.24	0.00	0.00	0.00	0.00
		Other Urban	0.40	0.33	0.00	0.00	0.00	0.00
		Non-Urban	1.17	0.96	0.01	0.00	0.01	0.02
	Renter	Urb>250000	1.66	1.17	0.01	0.01	0.01	0.01
		Other Urban	1.11	0.56	0.00	0.00	0.00	0.01
		Non-Urban	1.56	1.11	0.02	0.01	0.02	0.03
Black	Owner	Urb>250000	0.87	1.06	-0.06	-0.05	-0.06	-0.07
	Renter	Urb>250000	1.79	2.06	-0.13	-0.10	-0.13	-0.13
NBHIS	Owner	Urb>250000	2.48	2.17	-0.03	-0.01	-0.03	-0.02
	Renter	Urb>250000	3.49	4.55	-0.10	-0.03	-0.10	-0.10
		South:						
NHWO	Owner	Urb>250000	0.71	0.43	0.02	0.01	0.02	0.04
		Other Urban	0.42	0.35	0.02	0.01	0.02	0.04
		Non-Urban	0.69	0.56	0.03	0.02	0.04	0.05
	Renter	Urb>250000	1.52	0.97	0.03	0.03	0.04	0.06
		Other Urban	1.80	1.08	0.06	0.05	0.07	0.10
		Non-Urban	1.81	1.34	0.08	0.06	0.09	0.13
Black	Owner	Urb>250000	0.92	0.96	-0.08	-0.07	-0.08	-0.09
	Renter	Urb>250000	2.04	1.90	-0.16	-0.13	-0.17	-0.18
NBHIS	Owner	Urb>250000	0.92	0.48	-0.01	-0.01	-0.01	-0.01
	Renter	Urb>250000	2.82	1.07	-0.04	-0.02	-0.04	-0.04
		West:						
NHWO	Owner	Urb>250000	0.65	0.29	0.01	0.00	0.01	0.03
		Other Urban	0.58	0.49	0.02	0.01	0.02	0.05
		Non-Urban	0.69	0.58	0.02	0.01	0.02	0.06
	Renter	Urb>250000	1.67	0.91	0.02	0.01	0.02	0.07
		Other Urban	1.40	1.20	0.03	0.02	0.03	0.08
		Non-Urban	1.93	1.65	0.05	0.02	0.06	0.16
Black	Owner	Urb>250000	2.03	0.71	-0.04	-0.02	-0.04	-0.03
	Renter	Urb>250000	3.02	1.59	-0.08	-0.03	-0.08	-0.06
NBHIS	Owner	Urb>250000	0.90	0.39	-0.01	0.00	-0.01	0.00
	Renter	Urb>250000	1.96	1.38	-0.08	-0.05	-0.08	-0.06
		U.S.:						
Black	Owner	Other Urban	1.00	0.73	-0.06	-0.05	-0.06	-0.05
		Non-Urban	1.96	1.80	-0.16	-0.13	-0.16	-0.15
	Renter	Other Urban	1.23	0.70	-0.06	-0.04	-0.06	-0.05
		Non-Urban	5.68	6.55	-0.46	-0.35	-0.46	-0.41
NBHIS	Owner	Other Urban	1.68	1.44	-0.04	-0.04	-0.04	-0.02
		Non-Urban	2.75	1.67	0.00	-0.02	0.00	0.04
	Renter	Other Urban	2.90	2.43	-0.09	-0.06	-0.10	-0.08
		Non-Urban	6.09	6.34	-0.09	-0.11	-0.09	0.02
Asian	Owner		1.48*	0.60	0.01	0.01	0.01	0.02
Asian	Renter		2.70*	1.47	-0.04	0.01	-0.04	-0.02
PACIS	Owner		1.48*	2.50	-0.02	0.03	-0.02	0.02
PACIS	Renter		2.70*	4.73	-0.05	0.01	-0.05	0.01
American Indian Country AIR/TJSA/TDSA/ANVSA				6.39	-0.42	-0.13	-0.42	-0.42
American Indian Reservation AIR Only			5.25	3.24	0.00	0.00	0.00	0.00

[1]The simulated CVs presented in this table are in expectation and may vary based on the actual sample selected.

* Asian and Pacific Islanders were in the same major group in 1990. Asian and Pacific Islander Owners had a CV of 1.48%. Renters had a CV of 2.70%.

Simulated Person Sample Sizes[1]

Race/Eth.	Tenure	Geo. Group	1990 E-Sample	Prop. Allocation	Difference from Proportional Allocation: Optimal D/T Group	Optimal Minority 50% +	Emph. Minority	Keep All Minority
		North East:						
NHWO	Owner	Urb>250000	16,753	37,200	(3,600)	(2,000)	(4,100)	(6,100)
		Other Urban	11,196	12,400	(900)	(500)	(1,000)	(1,700)
		Non-Urban	13,512	25,000	(2,100)	(1,200)	(2,400)	(3,800)
	Renter	Urb>250000	7,163	19,000	(800)	(900)	(900)	(1,400)
		Other Urban	6,428	6,900	(300)	(200)	(300)	(500)
		Non-Urban	2,766	5,000	(300)	(200)	(400)	(700)
Black	Owner	Urb>250000	7,337	3,700	1,400	1,200	1,600	2,200
	Renter	Urb>250000	8,200	7,900	3,300	2,300	3,800	5,800
NBHIS	Owner	Urb>250000	1,251	1,800	300	100	300	500
	Renter	Urb>250000	3,905	6,100	2,300	1,200	2,700	4,400
		Midwest:						
NHWO	Owner	Urb>250000	15,778	35,400	(500)	(400)	(500)	(500)
		Other Urban	17,403	27,900	(600)	(300)		(800)
		Non-Urban	14,375	38,100	(900)	(600)	(1,000)	(1,500)
	Renter	Urb>250000	4,549	14,400	200	(100)	200	200
		Other Urban	7,174	13,800	(200)	(200)	(200)	(100)
		Non-Urban	3,146	8,200	(200)	(100)	(200)	(200)
Black	Owner	Urb>250000	8,226	4,600	700	500	700	900
	Renter	Urb>250000	8,607	6,500	1,000	800	1,000	1,300
NBHIS	Owner	Urb>250000	1,175	1,300	0	0	0	0
	Renter	Urb>250000	1,597	1,500	100	0	100	100
		South:						
NHWO	Owner	Urb>250000	18,920	41,200	(3,000)	(2,200)	(3,100)	(3,900)
		Other Urban	20,208	29,500	(2,700)	(2,000)	(2,800)	(3,500)
		Non-Urban	20,385	52,300	(4,700)	(3,600)	(4,900)	(6,200)
	Renter	Urb>250000	7,052	22,200	(1,000)	(1,200)	(1,000)	(1,200)
		Other Urban	11,057	15,300	(900)	(1,000)	(1,000)	(1,200)
		Non-Urban	4,255	10,600	(800)	(600)	(900)	(1,100)
Black	Owner	Urb>250000	11,195	8,200	2,200	2,000	2,300	3,400
	Renter	Urb>250000	9,020	11,700	3,000	2,500	3,200	4,300
NBHIS	Owner	Urb>250000	4,978	4,900	400	400	400	400
	Renter	Urb>250000	4,492	6,100	700	500	800	900
		West:						
NHWO	Owner	Urb>250000	14,528	38,900	(2,000)	(1,400)	(2,200)	(3,600)
		Other Urban	15,545	22,700	(1,000)	(600)	(1,100)	(1,500)
		Non-Urban	10,504	20,100	(900)	(300)	(1,000)	(1,800)
	Renter	Urb>250000	6,647	25,400	(700)	(700)	(800)	(900)
		Other Urban	8,912	12,700	(200)	(300)	(200)	200
		Non-Urban	3,600	6,100	(200)	0	(200)	(400)
Black	Owner	Urb>250000	2,385	2,400	400	300	400	500
	Renter	Urb>250000	2,353	4,300	500	200	600	800
NBHIS	Owner	Urb>250000	5,212	7,500	700	500	800	1,200
	Renter	Urb>250000	5,508	10,900	1,600	1,100	1,700	2,300
		U.S.:						
Black	Owner	Other Urban	7,066	6,400	1,600	1,400	1,700	2,100
		Non-Urban	2,662	7,100	2,100	1,700	2,200	2,500
	Renter	Other Urban	8,929	8,300	2,100	1,500	2,200	2,700
		Non-Urban	1,143	2,600	600	500	600	800
NBHIS	Owner	Other Urban	4,681	6,400	500	500	600	1,100
		Non-Urban	1,705	4,200	200	200	200	300
	Renter	Other Urban	4,591	5,800	700	400	700	1,300
		Non-Urban	1,044	2,400	100	100	100	200
Asian As an PACIS	Owner			11,900	400	(300)	400	400
	Renter			8,900	800	0	900	1,300
	Owner			800	100	0	100	100
ACIS	Renter			900	0	0	0	0
American Indian Country AIR/TJSA/TDSA/ANVSA				12,900	300	300	300	300
American Indian Reservation AIR Only			3,740	12,200	0	0	0	0

[1] The simulated sample sizes presented in this table are in expectation and may vary based on the actual sample selected.

Methodology for Obtaining Person Sample Sizes by Demographic/Tenure

For each cluster, we have a current estimate of housing units and estimates of the 1990 demographic composition of persons from the listing sample selection. There are some problems with trying to use this data directly to simulate sample size for different alternatives. First, our person estimates for 1990 don't necessarily correspond to current housing unit counts. We want to reflect population growth to the extent we can.

Second, when the estimated level of 1990 people in a cluster was different from the housing unit count, we were concerned with using the 1990 demographic distribution. For example, there is a cluster that has 403 HUs that had only 3 people (all Hispanic Renters) in 1990. Based on these types of clusters, we wanted the number of estimated people in the cluster to be related to our estimated housing unit count. Also, some sample clusters have many housing units but had no population in 1990.

A. Total Persons

The estimated number of people in each cluster was determined by converting the housing unit count to people. Using 1996 Census Bureau estimates, we first converted housing units to households and then to people by using the average number of people per household. This data is available separately by state.

$$\text{Est. People} = \text{HU} \times \frac{\text{Est. HH}}{\text{Est. HU}} \times \frac{\text{People}}{\text{per HH}}$$

It is important to note that this reflects some overall growth. However, it does not reflect differential growth rates of population subgroups. We have estimates for population subgroup changes form 1990 for larger areas like counties but we will not know the change for a specific cluster until after the Census.

B. Demographic Distribution

For each cluster, the estimated number of people in each of the 12 demographic/tenure groups is needed. The method for determining the demographic composition of a cluster is as follows.

If the cluster had 30 or more people in 1990, then the 1990 distribution of the cluster was used. If the cluster had fewer than 30 people, an average distribution was used instead. This approach reflects the variability of the cluster distributions while using an alternative for clusters whose 1990 distribution was based on a small sample size.

Average demographic distributions were computed on the cross-classification of the following variables:

- State
- Sampling Stratum
  - 1 = Small
  - 2 = Medium
  - 3 = Large
  - 4 = AIR
- Urbanicity
  - 1 = Urbanized Areas 250,000+
  - 2 = Other Urban
  - 3 = Non-Urban
- American Indian Country Indicator
  - 0 = Not AIC Land
  - 1 = AIR
  - 2 = AIC not AIR

The demographic distribution assigned to a cluster was based on the clusters affiliation to a cell of this cross-classificaiton.

C. Expected Sample Size

The expected sample size for a 1990 Major grouping can now be estimated based on alternative sample allocations. The sampling rates for the reduction subsampling strata will be specified by the allocation.

The sample size for the *i*th major group from the *j*th subsampling stratum in the *s*th state, in expectation, can be estimated by:

$$n_{s,j,i} = \sum_{k \in s,j} \frac{P_i}{TE_R}$$

where $n_{s,j,i}$ = expected sample size for the *i*th major group from the *j*th subsampling stratum in the *s*th state,
P_i = number of people in *k*th cluster who are in the *i*th major group,
TE_R = inverse of the sampling rate for the *j*th subsampling strata in the *s*th state.

Optimizing 1990 Major Groups

In order to generate an optimal allocation, the goal is to:

$$\text{Minimize} \sum_i \lambda_i \, CV_i^2$$

where i = 1 of 53 poststrata (reflecting the split of Asians and Pacific Islanders) and
λ_i = influence of the ith major group in the optimization,

based on the constraint that for each state, s, :

$$\sum_j \sum_{k \in j} \frac{b_{s,j}}{B_{s,j}} HU_k = \begin{array}{c}\text{A.C.E. State Interview} \\ \text{HU Sample Size}\end{array}$$

where j = A.C.E. reduction sampling strata,
k = listing sample clusters in jth stratum,
$b_{s,j}$ = the number of A.C.E. reduction clusters in the jth stratum to be selected (the quantity to be optimized.),
$B_{s,j}$ = the total number of A.C.E. listing sample clusters in the jth stratum and
HU_k = the number of interview housing unit counts in the kth cluster in the jth stratum (reflects large block cluster subsampling).

Lagrange multiplier can be used to derive the following constrained optimal allocation for the jth subsampling stratum in a state.

$$b_{s,j} = \frac{S \times B_{s,j} \sqrt{\dfrac{\sum_{k \in s,j} P_k \, Wgt_k^2 \sum_{i \in k} \dfrac{\lambda_i \sigma_i^2 prop_{k,i}}{DSE_i^2}}{\sum_{k \in s,j} HU_k}}}{\sum_j \left[\left(\sqrt{\dfrac{\sum_{k \in s,j} P_k \, Wgt_k^2 \sum_{i \in k} \dfrac{\lambda_i \sigma_i^2 prop_{k,i}}{DSE_i^2}}{\sum_{k \in s,j} HU_k}} \right) \times \sum_{k \in s,j} HU_k \right]}$$

where Wgt_k is the listing sample cluster weight adjusted for large block cluster subsampling,
S is the A.C.E. state housing unit interview sample size,
P_k is the estimated number of people in the *k*th cluster,
σ_i^2 is the unit variance for the *i*th major group,
DSE_i^2 is the squared Dual System Estimate for *i*th major group for 1990,
HU_k is the number of housing units in the *k*th cluster and
$prop_{k,i}$ is the proportion of people in the *k*th cluster who fall into the *i*th major group.

Reliability Estimation Methodology

The following gives more details of how CVs are simulated for the major groups once sample size and sampling rates are determined for a particular allocation.

A. Obtain Variance Component Independent of Sample Size and Weights.

The 1990 PES design had differential weights for the sampling strata. For each major group, this step factored out the effect of sample size and the differential weighting of the 1990 PES from the DSE variance.

This variance component of the DSE in a major group, σ_i^2, (i = 1 to 51) was calculated from the following formula:

$$\sigma_i^2 = \frac{\text{Var } (DSE_i)}{\sum_{j=1}^{n_i} w_{i,j}^2}$$

where Var (DSE_i) = 1990 PES Variance estimated by the Jackknife methodology,
$w_{i,j}$ = the inverse probability of selection of the *j*th E-sample person in the *i*th major group,
n_i = the number of E-sample people in the *i*th major group in the 1990 PES.

B. Estimate Major Group CVs for Alternative Sample Designs

The DSE variance for a major group is a function of the unit variance component, the sample size and the weights. The amount of sample and the weights in each major group will change from the 1990 PES to the allocation alternative design. We assumed the major group unit variance components computed above were the same as in 1990.

For this research, separate estimates of Asians and Hawaiian and Pacific Islanders are made. This required the assumption that unit variance component for both the Asian and the Hawaiian and Pacific Islander Owners were equal to the unit variance component calculated for the Asian and Pacific Islander Owners. A similar assumption is required for Renters.

An estimate of the DSE variance for a major group was calculated based on the sample size and weights for an allocation alternative. The estimated sample size is an expectation based on the alternative sample design and the estimated cluster characteristics. The actual sample selected

may differ from these estimates. There are other differences between the 1990 PES and the A.C.E. that need to be accounted for in our CVs: 1) the change in surrounding block search methodology and 2) the effect of small block cluster weighting. To compensate for the former change a scalar of 1.56 was applied, while a scalar of 1.14 was used for the latter. These adjustments may change based on further research.

$$\mathrm{Var}\,(DSE_i^*) \;=\; 1.56 \times 1.14 \times \sigma_i^2 \sum_s \sum_{k \in j} n^*_{s,j,k,i}\; W^{*2}_{s,j}$$

where $n_{s,j,k,i}^*$ = expected number of people in the *k*th cluster from the *j*th subsampling stratum in state *s* who are in the *i*th major group,
$W_{s,j}^*$ = the inverse probability of selection of the block cluster in the *j*th subsampling stratum in state *s* for the alternative sample design.

The estimated CV for each major group (in percent) is then calculated as follows:

$$CV_i^* = \left(\frac{\sqrt{\mathrm{Var}\,(DSE_i^*)}}{DSE_i} \right) \times 100$$

where DSE_i = the 1990 DSE for the major group.